KE LAB ASSIGNMENT – 3

By:-
Anshuman Singh
Roll: 187107
B.Tech CSE 3rd year
Section A

Q1. Missing data problem. Find the missing data elements and replace them with the avg of that attribute.

Sol:

```
#include<bits/stdc++.h>
using namespace std;
int main()
{
freopen("dataset.txt","r",stdin);
freopen("output.txt","w",stdout);
string dat,foravg;
int avg=70,sum=0,count=0;

while(getline(cin,dat)){
int i=0;
while(dat[i]>='0'&&dat[i]<='9'){
i++;
}
if(i+1==dat.length()){
avg=sum/count;
sum=sum+avg;
count++;
cout<<dat<<avg;
}
else{
cout<<dat;
int t;
stringstream stream(dat.substr(i+1));
stream>>t;
//cout<<" "<<dat.substr(i+1);
sum+=t;
count++;
}
cout<<"\n";
}
return 0;
}
```

Input:
001 67
002 82
003

004 75
005 91
006 59
007
008 86
009 62

Output:
001 67
002 82
003 74
004 75
005 91
006 59
007 74
008 86
009 62

Q2. Missing value problem. Find the missing data elements and replace them with a global constant.

Sol:

```
#include<bits/stdc++.h>
using namespace std;
int main()
{
freopen("dataset.txt","r",stdin);
freopen("output.txt","w",stdout);
string dat,foravg;
int avg=70,sum=0,count=0;

while(getline(cin,dat)){
int i=0;
while(dat[i]>='0'&&dat[i]<='9'){
i++;
}
if(i+1==dat.length()){
cout<<dat<<avg;
}
else{
cout<<dat;
}
cout<<"\n";
}
return 0;
}
```

Input:
001 67
002 82
003
004 75

005 91
006 59
007
008 86
009 62

Output:
001 67
002 82
003 70
004 75
005 91
006 59
007 70
008 86
009 62

Q3. Noisy data problem. Use binning technique to remove noise data. Use smooth by bin means method.

Sol:

```
#include<bits/stdc++.h>
using namespace std;
int main()
{
freopen("input.txt","r",stdin);
freopen("output.txt","w",stdout);
string dat;
vector<int> val;
int num_bins=3;

while(cin>>dat){
int t;
stringstream stream(dat);
stream>>t;
val.push_back(t);
}
sort(val.begin(),val.end());
int width = val.size()/num_bins;
for(int i=0;i<val.size();i++){
int avg,sum=0,c=0;
for(int j=i;j<i+width;j++){
sum+=val[j];
c++;
}
avg=sum/c;
for(int j=i;j<i+width;j++){
val[j]=avg;
}
i=i+width-1;
```

```
}
for(int i=0;i<val.size();i++){
cout<<val[i]<<" ";
}
return 0;
}
```

Input:
1 5 2 7 23 45 13 23 57 52 24 17

Output:
3 3 3 3 19 19 19 19 44 44 44 44

Q4. Noisy data problem. Use binning technique to remove noise data. Use smooth by bin boundaries method.

Sol:

```
#include<bits/stdc++.h>
using namespace std;
int main()
{
freopen("input.txt","r",stdin);
freopen("output.txt","w",stdout);
string dat;
vector<int> val;
int num_bins=3;

while(cin>>dat){
int t;
stringstream stream(dat);
stream>>t;
val.push_back(t);
}
sort(val.begin(),val.end());
int width = val.size()/num_bins;
for(int i=0;i<val.size();i++){
for(int j=i;j<i+width;j++){
if(abs(val[j]-val[i])<abs(val[j]-val[i+width-1])){
val[j]=val[i];
}
else{
val[j]=val[i+width-1];
}
}
i=i+width-1;
}
for(int i=0;i<val.size();i++){
cout<<val[i]<<" ";
}
return 0;
```

```
}
```

Input:
1 5 2 7 23 45 13 23 57 52 24 17

Output:
1 1 7 7 13 13 23 23 24 57 57 57